Space-Frequency-Interpolated Radio Map

Koya Sato, *Member, IEEE*, Katsuya Suto, *Member, IEEE*, Kei Inage, *Member, IEEE*, Koichi Adachi, *Senior Member, IEEE*, and Takeo Fujii, *Member, IEEE*

***Abstract*—This paper presents a novel method for radio map construction that simultaneously interpolates the received signal power values over space and frequency domains. Radio maps can be used to improve spectrum management and for localization systems, which are related to wireless systems in general such as cellular systems, Internet of Things (IoT) networks, and vehicular communications. Researchers have shown that crowdsourcing using spatial interpolation techniques can be used to accurately construct a radio map to improve these applications; however, because of the limitation in the spatial domain, conventional methods can build a radio map for only those frequencies at which sensing can be performed. Our proposed method focuses on the fact that the shadowing values show strong correlation over a wide range of frequency domains. The main idea is to treat the shadowing values obtained over various frequencies as the ones over target frequency. Using the actual datasets obtained over 870, 2115, and 3500 MHz in a cellular system, we show that the proposed method can accurately generate a radio map, even if no (or few) datasets are available in the target frequency.**

***Index Terms*—Correlation, crowdsourcing, interpolation, radio map, radio propagation, spatial statistics.**

I. Introduction

A RADIO map is a tool for visualizing the spatial distribution of communication quality, such as the average received signal power [2]. Because of its simple definition, it has been widely applied over the last decade in fields such as fingerprint-based localization for Internet of things (IoT) devices [3]–[9], resource management in cellular networks [10]–[13], and spectrum sharing over white spaces [14]–[16]. Generally, an accurate radio map can enhance the performance of such a radio-map-assisted system (e.g., spectral efficiency in spectrum sharing [2] and localization accuracy in the fingerprint-based localization). This fact has motivated many researchers to study the accurate construction of radio maps.

Manuscript received October 6, 2020; revised November 26, 2020 and December 25, 2020; accepted December 29, 2020. Date of publication January 8, 2021; date of current version February 12, 2021. This work was supported by the Ministry of Internal Affairs and Communications of Japan under Grant JPJ000254. An early work related to this paper was presented at IEEE International Conference on Artificial Intelligence in Information and Communication (ICAIIC) 2020, [1]. The review of this article was coordinated by Dr. Guan Gui. *(Corresponding author: Koya Sato.)*

Koya Sato is with the Department of Electrical Engineering, Tokyo University of Science, Tokyo 125-8585, Japan (e-mail: k_sato@ieee.org).

Katsuya Suto is with the Graduate School of Informatics and Engineering, The University of Electro-Communications, Tokyo 182-8585, Japan (e-mail: k.suto@uec.ac.jp).

Kei Inage is with the Tokyo Metropolitan College of Industrial Technology, Tokyo 140-0011, Japan (e-mail: inage@metro-cit.ac.jp).

Koichi Adachi and Takeo Fujii are with the Advanced Wireless and Communication Research Center (AWCC), The University of Electro-Communications, Tokyo 182-8585, Japan (e-mail: adachi@awcc.uec.ac.jp; fujii@awcc.uec.ac.jp).

Digital Object Identifier 10.1109/TVT.2021.3049894

A. Motivation

Methods for the construction of radio maps can be mainly categorized into an empirical path-loss-model-based approach or a crowdsourcing-based approach [17]. Although the former method can easily construct radio maps, its accuracy is poor because of uncertainties in the local variation of the radio propagation, such as shadowing [18]. In contrast, the crowdsourcing-assisted radio map using spatial interpolation [19] can provide highly accurate radio maps. In this method, a cloud server first collects the received signal values related to the received location from massive mobile terminals (e.g., smartphones or vehicles). Next, the radio map is constructed by spatially estimating the missing information via a regression technique, such as Kriging (or Gaussian process regression) [19] and machine learning [20], [21]. However, these related techniques can interpolate the radio map only over the space domain; that is, the frequency bands for which a radio map can be constructed are limited to frequencies available to the target transmitter at the time of the measurement campaign.

In recent years, spectrum use in various bands has been restructured; for example, bands that were previously used for television and radar are being opened for spectrum-sharing with secondary systems [22]. However, in general, radio map construction requires a massive amount of data at the frequency of interest because conventional techniques for radio map construction can only be used for interpolation over the spatial domain. When a new frequency is opened for a radio-map-assisted system of interest, conventional radio map construction methods would need a site survey of the corresponding frequency. This would incur additional considerable sensing costs.

Motivated by this background, we propose a technique that interpolates the radio map not only over the spatial domain but also the frequency domain.

B. Related Works

The average received power consists of path loss and shadowing. Therefore, it is important to estimate these two factors accurately for highly accurate radio maps. Some techniques can estimate the radio propagation over both spatial and frequency domains jointly; however, these techniques can only estimate the path loss.

Empirical path loss models (e.g., Okumura-Hata model [23], [24], irregular terrain models [25], and COST-231 [26]) are traditional enablers for the above motivation. A suitable choice of the path loss model will obtain a roughly accurate radio map. However, if path loss indexes differ between the application

environment and the selected model, the estimation accuracy significantly deteriorates. Although we can compensate for such a difference by fitting the path loss model to a few measurement data [27], [28], the estimation accuracy is still limited because such a method cannot estimate the local variation owing to shadowing effects [18].

Recent works have shown that feedforward neural networks (FFNNs) with crowdsourcing can obtain an accurate path loss [29]–[34]. Ayadi *et al.* proposed an FFNN-based path loss modeling method in which the input layer requires some features of the path loss, such as communication distance and frequency [29]. They showed that this FFNN can predict path loss over multibands in the ultra-high frequency (UHF). Additionally, similar results have been reported in other situations, such as aircraft cabin environments [30] and very high frequency (VHF) bands [34]. Designs of hyper-parameters are discussed by Sotiroudis *et al.* in [33] and Popoola *et al.* in [35]. Such an FFNN-based method is advantageous in terms of applicability; however, FFNN tends to regard shadowing as Gaussian noise to avoid overfitting. This fact still limits the radio propagation estimation accuracy, although the FFNN can model anisotropy in the path loss curves [36].

As described above, the related techniques cannot estimate shadowing, and this fact limits the estimation accuracy. For more accurate space-frequency interpolation, it is important to estimate both the path loss and shadowing, similar to spatial interpolation using Kriging.

Some researchers in the field of radio propagation have revealed that shadowing is highly correlated in not only the spatial domain but also the frequency domain [37]–[40]. For example, the authors in [38] measured the frequency correlation between the uplink and downlink signals in the GSM1900 system (i.e., 1900 MHz with a frequency domain duplex (FDD) system). The results showed that the correlations were roughly 0.90 in suburban areas and 0.75 in rural areas. Additionally, [39] evaluated the frequency correlation between 900, 1800, and 2100 MHz bands in urban environments; they found that the correlation between the 900 and 1800 MHz bands is 0.84, and the correlation between the 900 and 2100 MHz bands is 0.79. This empirical property naturally raises a simple idea: *the radio map could be interpolated over spatial and frequency domains accurately*.

C. Objective and Major Contributions of This Paper

In this study, we propose a joint space-frequency interpolation technique for radio map construction. The main idea is to treat shadowing values over multiple frequency bands as the ones over the target frequency. The proposed method first performs regression analysis over distance and frequency domains for path loss modeling. After shadowing at the target frequency is interpolated over the frequency domain, shadowing factors at the target band are spatially interpolated via ordinary Kriging. Finally, for better accuracy, this method can calibrate the path loss characteristics over the target band. This option can be applied if we obtain few datasets over the target frequency band, which can be employed to modify the constant factor in the path loss curve.

Actual datasets have been obtained over three cellular bands to evaluate the performance of the proposed method—870, 2115, and 3500 MHz. First, experimental results demonstrate that the proposed method can accurately interpolate the radio map of 2115 MHz from datasets obtained over 870 and 3500 MHz and a few measurements over 2115 MHz. Additionally, in this study, we also performed frequency extrapolation, wherein the radio map of 3500 MHz is estimated from datasets in 870 and 2115 MHz. Thus, the proposed method enables us to construct a radio map over the frequency with no (or a few) measurement data.

The main contributions of this study are listed as follows:

- We extend the Kriging-based radio map construction from the spatial interpolation to space-frequency interpolation, which enables us to obtain an accurate radio map on a frequency band where no (or a few) datasets are available. This extension can be realized simply by interposing a linear interpolation of shadowing values on the frequency axis in advance of the conventional spatial interpolation.
- A large-scale measurement campaign over cellular bands reveals that the proposed method outperforms the conventional methods (neural network-based method and fine-tuned path loss) in terms of the interpolation accuracy.
- Furthermore, it is revealed that the proposed approach enables not only frequency interpolation but also extrapolation.

Note that this paper is an extended version of work we presented earlier [1]. The major new contribution of the work presented in this paper is to introduce a re-estimation algorithm for path loss modeling using a few datasets at the frequency of interest. This algorithm significantly improves the accuracy of the radio map. Additionally, in contrast to our previous study, which was simulation-based [1], this study reveals the practical performance of the proposed method using an actual dataset of a cellular system.

D. Organization

The remainder of this paper is organized as follows. In Sect. II, the objective and system model of the radio map construction are defined. Next, Sect. III reviews the conventional discussions on path loss and shadowing in the field of radio propagation; in Sect. IV, we present the proposed space-frequency interpolation, which was designed based on the reviews in Sect. III. Sect. V explains the details of the measurement campaign; the performances of the proposed method are discussed in Sect. VI. Finally, Sect. VII concludes this study.

II. Objective and System Model

This section presents the objectives of the study and the system model. This study assumes a situation wherein a transmitter provides its own service with multi-bands, and the radio map is utilized to estimate the radio propagation characteristics of this transmitter. Specifically, we aim to reduce the sensing cost for radio map construction, i.e., the number of datasets on the frequency band of interest via space-frequency interpolation. Some of the major applications of our proposed method include

cell planning in multi-band cellular systems and a site survey on frequencies that have been recently made available to existing wireless systems (e.g., unlicensed long-term evolution (U-LTE) [41]).

This study defines the radio map as a map that has two-dimensional coordinate information of the average received signal power from the target transmitter. We assume that radio maps are stored in the cloud database. First, the database collects the received signal power on the mobile terminal, which is related to the received location and the carrier frequency. Afterwards, a single radio map is constructed for each frequency. Therefore, we consider the situation wherein a fixed transmitter is located on the coordinates $\boldsymbol{x}_{\mathrm{Tx}}$ and can transmit signals on $M+1$ frequencies, $\{f_0, f_1, f_2, \ldots, f_M\}$, where f_0 is the target frequency and f_k $(k = 1, 2, \ldots, M)$ is the measurable frequency used for interpolating the radio map at frequency f_0. Note that the magnitude relationship of frequencies does not depend on index k and can be in any order.

In a typical radio map construction, multiple nodes first measure the signal strength from the transmitter and report the data to the database. Thereafter, all instantaneous values are averaged for each frequency and grid (from several meters to several hundred meters) to alleviate the effect of multipath fading.[1] Throughout this study, we refer to the received signal power as the averaged value.

Consider $P(\boldsymbol{x}, f)$ [dBm] as the received signal power at a grid related with the coordinate $\boldsymbol{x}$ and a carrier frequency f [MHz]. The database has N received signal power values. Each value can be obtained on a different grid and frequency, and we express the number of received signal values on f_k as N_k $(k = 1, 2, \ldots, M,$ and $\sum_{k=1}^{M} N_k = N)$. Therefore, the database stores the dataset over f_k, which is given as follows:

$$\boldsymbol{y}_{f_k} = \left(P(\boldsymbol{x}_{k,1}, f_k), P(\boldsymbol{x}_{k,2}, f_k), \ldots, P(\boldsymbol{x}_{k,N_k}, f_k)\right)^{\top}, \quad (1)$$

where $(\cdot)^{\top}$ represents the operation of transpose, and $\boldsymbol{x}_{k,j}$ is the j-th measurement location on the frequency f_k. The objective of this paper is to interpolate the received signal value $P(\boldsymbol{x}_0, f_0)$ using the concatenation of vectors $\boldsymbol{y} = (\boldsymbol{y}_{f_1}, \boldsymbol{y}_{f_2}, \ldots, \boldsymbol{y}_{f_M})$. In other words, we aim to estimate the received signal at an arbitrary grid related to the coordinate $\boldsymbol{x}_0$ over the partial or non-measured frequency f_0.

Furthermore, this study considers two scenarios for measurement that concern the target frequency: (i) no measurement data for f_0 $(N_0 = 0)$ and (ii) an insignificant amount of measurement data for f_0 $(0 < N_0 \ll N$, and these data form a dataset vector $\boldsymbol{y}_{f_0})$. Based on this consideration, we investigate the effect of an optional dataset $\boldsymbol{y}_{f_0}$ on the accuracy of radio map construction.

III. Review of Radio Propagation Characteristics

Next, we review the detailed characteristics of $P(\boldsymbol{x}, f)$ for designing the joint space-frequency interpolation. If the effect of multipath fading is completely removed, the received signal power at an arbitrary location $\boldsymbol{x}$ and a carrier frequency f can be expressed as follows:

$$P(\boldsymbol{x}, f) = \overline{P}(\boldsymbol{x}, f) + W(\boldsymbol{x}, f), \quad (2)$$

where $\overline{P}(\boldsymbol{x}, f)$ [dBm] is the large-scale trend that contains transmission power, antenna gain, effect of frequency, and path loss. Numerous traditional empirical path loss models (such as the Okumura-Hata model [27]) indicate that the path loss increases in proportion to $\log_{10} f$ and $\log_{10} ||\boldsymbol{x}_{\mathrm{Tx}} - \boldsymbol{x}||$ (where $||\cdot||$ represents Euclidean distance). Based on this discussion, $\overline{P}(\boldsymbol{x}, f)$ can be modeled as

$$\begin{aligned}\overline{P}(\boldsymbol{x}, f) &= P_{\mathrm{Tx}} - L_{\mathrm{C}} - \eta_f \log_{10} f - 10\eta_d \log_{10} d \\ &\triangleq P_{\mathrm{C}} - \eta_f \log_{10} f - 10\eta_d \log_{10} d, \end{aligned} \quad (3)$$

where P_{Tx} [dBm] is the transmission power, L_{C} [dB] is the constant propagation loss, $d \triangleq ||\boldsymbol{x}_{\mathrm{Tx}} - \boldsymbol{x}||$ [m], $P_{\mathrm{C}} \triangleq P_{\mathrm{Tx}} - L_{\mathrm{C}}$ [dBm], and η_f and η_d are path loss indexes in terms of frequency and distance, respectively. Additionally, $W(\boldsymbol{x}, f)$ [dB] is the log-normal shadowing with zero mean and standard deviation σ [dB]. Based on its typical definition (e.g., [42]), this study defines shadowing as a stochastic fluctuation of received signal values caused by obstacles between the transmitter and receiver that attenuate signal power through absorption, reflection, scattering, and diffraction. The shadowing fluctuates the received signal values over distances proportional to the length of the obstructing object (10-100 m in outdoor environments). It can be extracted by taking the logarithmic difference between the received signal power values averaged per few meters and the estimated curve of $\overline{P}(\boldsymbol{x}, f)$.

It is well known that $W(\boldsymbol{x}, f)$ shows correlations over the spatial domain [43]. For two shadowing values with the same frequencies and different locations, the spatial correlation of shadowing is given by

$$\rho\left(W(\boldsymbol{x}_i, f), W(\boldsymbol{x}_j, f)\right) = \exp\left(-\frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|}{d_{\mathrm{cor}}} \ln 2\right), \quad (4)$$

where d_{cor} [m] is the correlation distance [43].

Finally, we consider the situation wherein two shadowing values are obtained over the same location $\boldsymbol{x}$ but over different frequencies. Some studies have reported that shadowing values also show correlation over the frequency domain, even if the frequency difference $|f_k - f_l|$ exceeds hundreds of MHz. Table I summarizes the main timeline of discussions on the frequency correlation of shadowing, which mainly focuses on outdoor environments. For example, Mogensen [37] measured shadowing correlation between 955 MHz and 1845 MHz in an urban area and found that the correlation was at least 0.85. Additionally, Perahia [38], Laethem [39], and Sato [40] reported similar results from measurement campaigns in outdoor environments over UHF bands or thereabouts. These discussions imply that the frequency correlation of shadowing is above 0.7 in outdoor environments over UHF bands or thereabouts.

[1] Multipath fading has the spatial correlation. However, its correlation distance is sufficiently shorter than the shadowing, and it shows the time-variant property, unlike the shadowing. If the effect of multipath fading remains in the dataset, its value propagates to the interpolation result as an uncertainty factor in Kriging-based interpolation; that is, the accuracy of the radio map will deteriorate if we do not eliminate the effect of multipath fading in advance.

TABLE I
RELATED WORKS ON FREQUENCY CORRELATION OF SHADOWING

Year	Authors	Frequencies	Environment	Main results
1991	Mogensen *et al.* [37]	955 MHz and 1845 MHz	Urban	Results from dual-frequency measurements performed in urban areas at 955 MHz and 1845 MHz were presented. They found the shadowing correlation to be at least 0.85 for various transmission locations.
2001	Perahia *et al.* [38]	1900 MHz (uplink and downlink in an FDD system)	Suburban and rural	They measured the frequency correlation between the uplink and downlink signals in the GSM1900 system (i.e., 1900 MHz with an FDD system). The results showed that the correlations were roughly 0.90 in suburban areas and 0.75 in rural areas.
2012	Laethem *et al.* [39]	900 MHz, 1800 MHz, and 2100 MHz	Urban	They performed their evaluation in an urban area. The frequency correlation between 900 MHz and 1800 MHz is 0.84, and the correlation between 900 and 2100 MHz is 0.79.
2018	Sato *et al.* [40]	521-551MHz	Suburban	They evaluated the frequency correlation of shadowing in television bands (6 channels in the range 521-551 MHz) in a suburban area. The shadowing correlation was greater than 0.90 for all frequency pairs.
2021	(This work)	870 MHz, 2115 MHz, and 3500 MHz	Urban	We performed a measurement campaign over three bands (870 MHz, 2115 MHz, and 3500 MHz) in an urban area. The measurement results revealed that the frequency correlation was greater than 0.90 for all frequency pairs (see Sect. V-B and Table III).

Based on this information, we can consider that the correlation between $W(\boldsymbol{x}, f_k)$ and $W(\boldsymbol{x}, f_l)$, i.e., $\rho(W(\boldsymbol{x}, f_k), W(\boldsymbol{x}, f_l))$ takes approximately any value in the range of 0.7–1.0.[2]

IV. KRIGING-BASED SPACE-FREQUENCY INTERPOLATION

This section presents the proposed method that expands on ordinary Kriging [19] from the spatial interpolation to the space-frequency interpolation. Usually, Kriging is performed for spatial interpolation over one frequency (i.e., only $\boldsymbol{y}_{f_0}$ is available with a sufficient number of datasets, and $\boldsymbol{y}$ is empty). Meanwhile, as reviewed in Sect. II, shadowing exhibits a strong frequency correlation. Focusing on this characteristic, we can accurately construct a radio map at f_0 using dataset $\boldsymbol{y}$, which contains several received signal power values over multiple frequencies.

The main principle is to treat the shadowing values obtained over various frequencies as the ones over the target frequency. In other words, the frequency correlations of the shadowing between different frequencies are treated as one. Thereafter, the weighted averaging based on Kriging is applied to these values *without considering the frequency correlation* to perform spatial interpolation.

Fig. 1 illustrates the flowchart of the proposed method. The proposed method estimates $\overline{P}(\boldsymbol{x}_0, f_0)$ and $W(\boldsymbol{x}_0, f_0)$ separately and is subdivided into the following five steps: (i) path loss modeling (with re-modeling if $N_0 > 0$), (ii) shadowing extraction and frequency interpolation, (iii) semivariogram modeling, (iv) shadowing estimation using ordinary Kriging, and (v) estimation of the received signal power. We discuss these steps as follows:

A. Path Loss Modeling Using $\boldsymbol{y}$

Ordinary Kriging achieves optimal interpolation for the Gaussian process, which follows a unique model wherein both the expected value and semivariogram of the random variable have spatial stationarity [19]. As the expected value of the received signal power exhibits a spatial trend because of the path loss, we first need to estimate the path loss and remove it from the dataset. To extract the shadowing, we estimate P_{C}, η_f, and η_d by fitting Eq. (3) into dataset $\boldsymbol{y}$ using ordinary least squares (OLS). Using OLS, the parameter vector $\boldsymbol{\theta} = (P_{\mathrm{C}}, \eta_f, \eta_d)^\top$ can be estimated as follows:

$$\boldsymbol{\theta} = \left(\boldsymbol{A}^\top \boldsymbol{A}\right)^{-1} \boldsymbol{A}^\top \boldsymbol{y}, \tag{5}$$

where

$$\boldsymbol{A} = \begin{pmatrix} 1 & -\log_{10} f_1 & -10\log_{10}||\boldsymbol{x}_{\mathrm{Tx}} - \boldsymbol{x}_{1,1}|| \\ 1 & -\log_{10} f_1 & -10\log_{10}||\boldsymbol{x}_{\mathrm{Tx}} - \boldsymbol{x}_{1,2}|| \\ \vdots & \vdots & \vdots \\ 1 & -\log_{10} f_1 & -10\log_{10}||\boldsymbol{x}_{\mathrm{Tx}} - \boldsymbol{x}_{1,N_1}|| \\ 1 & -\log_{10} f_2 & -10\log_{10}||\boldsymbol{x}_{\mathrm{Tx}} - \boldsymbol{x}_{2,1}|| \\ 1 & -\log_{10} f_2 & -10\log_{10}||\boldsymbol{x}_{\mathrm{Tx}} - \boldsymbol{x}_{2,2}|| \\ \vdots & \vdots & \vdots \\ 1 & -\log_{10} f_M & -10\log_{10}||\boldsymbol{x}_{\mathrm{Tx}} - \boldsymbol{x}_{M,N_M}|| \end{pmatrix}.$$

Thereafter, the estimated value of $\overline{P}(\boldsymbol{x}, f)$ is derived by

$$\overline{P}_{\mathrm{est}}(\boldsymbol{x}, f) = \hat{P}_{\mathrm{C}} - \hat{\eta}_f \log_{10} f - 10\hat{\eta}_d \log_{10}||\boldsymbol{x}_{\mathrm{Tx}} - \boldsymbol{x}||, \tag{6}$$

where $\hat{P}_{\mathrm{C}}$, $\hat{\eta}_f$, and $\hat{\eta}_d$ are estimated values for P_{C}, η_f, and η_d, respectively.

B. Extraction and Frequency Interpolation of Shadowing

The shadowing quantity of each datapoint in $\boldsymbol{y}$ is extracted using

$$\hat{W}(\boldsymbol{x}_{k,i}, f_k) = P(\boldsymbol{x}_{k,i}, f_k) - \overline{P}_{\mathrm{est}}(\boldsymbol{x}_{k,i}, f_k). \tag{7}$$

Next, shadowing $\hat{W}(\boldsymbol{x}_{k,i}, f_k)$ $(k = 1, 2, \ldots, M)$ is treated as the one at $\boldsymbol{x}_{0,i}$ and f_0, i.e., $\hat{W}(\boldsymbol{x}_{0,i}, f_0) = \hat{W}(\boldsymbol{x}_{0,i}, f_k)$. Note that because the radio map manages the received power values over spatially quantized grids, multiple frequencies may be present at

[2]Unfortunately, a general formula for the frequency correlation has not been developed currently; thus, we cannot express this characteristic using one equation. Establishing this model will be an important issue in the field of radio propagation.

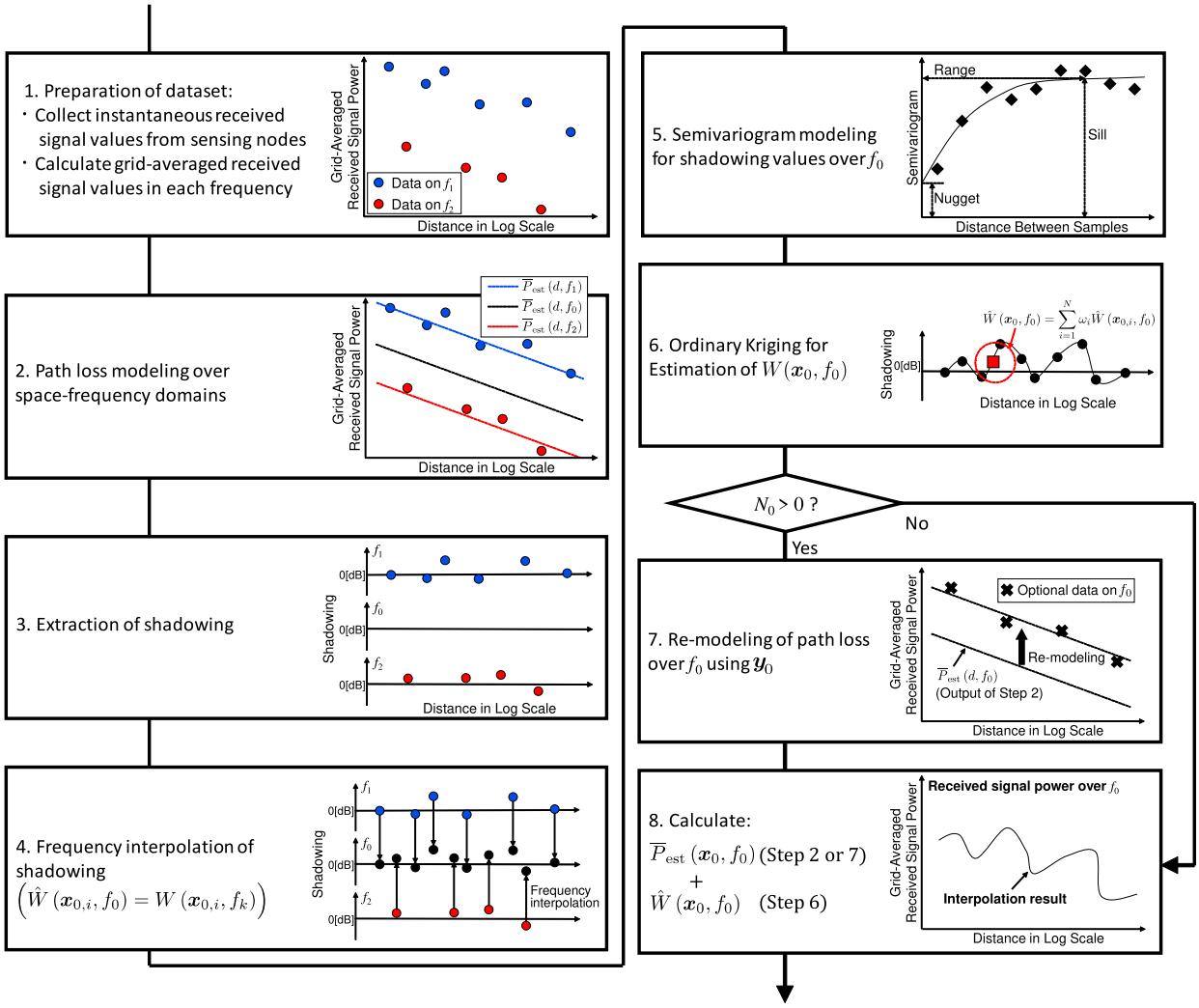


Fig. 1. Flowchart of the proposed radio map construction method (one-dimensional case with $M = 2$).

the same location. For this case, we treat the logarithmic average of these shadowing values as $\hat{W}(\boldsymbol{x}_{0,i}, f_0)$.

C. Semivariogram Modeling

Next, we model the structure of the spatial correlation when f_0 using a semivariogram. A semivariogram γ is defined as

$$\gamma(d_{i,j}) = \frac{1}{2}\mathrm{Var}[W(\boldsymbol{x}_{0,i}, f_0) - W(\boldsymbol{x}_{0,j}, f_0)], \tag{8}$$

where $d_{i,j} \triangleq ||\boldsymbol{x}_{0,i} - \boldsymbol{x}_{0,j}||$, and $\mathrm{Var}[\cdot]$ is the variance of the random variable. The semivariogram can be modeled using binning [19]. In this method, empirical semivariograms are generated from $\hat{W}(\boldsymbol{x}_i, f_0)$ obtained in Sect. IV-B. Subsequently, a theoretical semivariogram is fitted to these generated semivariograms. There are some theoretical semivariogram models (e.g., Gaussian, exponential, and spherical), and we have to choose a suitable model based on the spatial correlation property of the dataset. Here, as indicated in Eq. (4), the spatial correlation of shadowing follows the exponential decay model. Considering this background, we use the exponential semivariogram model, which is defined as

$$\gamma(d_{i,j}) = \alpha_n^2 + \alpha_s^2 \left\{1 - \exp\left(\frac{-d_{i,j}}{\alpha_r}\right)\right\}, \tag{9}$$

where α_n^2, α_s^2, and α_r are *nugget*, *sill*, and *range*, respectively; we can obtain these parameters by fitting Eq. (9) to the empirical semivariogram values via non-linear least squares.

D. Estimation of Shadowing Via Ordinary Kriging

Ordinary Kriging interpolates the target random variable by taking the weighted average of the datasets. In our context, this can be regarded as the estimation of $W(\boldsymbol{x}_0, f_0)$ from $\hat{W}(\boldsymbol{x}_i, f_0)$; this can be written as

$$\begin{aligned}\hat{W}(\boldsymbol{x}_0, f_0) &= \sum_{i=1}^{N} \omega_i \hat{W}(\boldsymbol{x}_{0,i}, f_0) \\ &= \sum_{k=1}^{M} \sum_{j=1}^{N_k} \omega_{k,j} \left(P(\boldsymbol{x}_{k,j}, f_k) - \overline{P}_{\rm est}(\boldsymbol{x}_{k,j}, f_k)\right),\end{aligned} \tag{10}$$

where ω_i and $\omega_{k,j}$ are the weight factors satisfying $\sum_{i=1}^{N}\omega_i = \sum_{k=1}^{M}\sum_{j=1}^{N_k}\omega_{k,j} = 1$. Note that later discussions are based on Eq. (10) for mathematical simplicity. Kriging optimizes ω_i such that the variance of the estimation error $\sigma_{\mathrm{e}}^2 = \mathrm{Var}[\hat{W}(\boldsymbol{x}_0, f_0) - W(\boldsymbol{x}_0, f_0)]$ is minimized. Using the Lagrange multiplier method, the objective function can be written as follows:

$$\phi(\omega_i, \mu) = \sigma_{\mathrm{e}}^2 - 2\mu\left(\sum_{i=1}^{N}\omega_i - 1\right), \tag{11}$$

where μ is the Lagrange multiplier. Here, σ_{e}^2 can be given by (see [19]):

$$\sigma_{\mathrm{e}}^2 = -\gamma(d_{0,0}) - \sum_{i=1}^{N}\sum_{j=1}^{N}\omega_i\omega_j\gamma(d_{i,j}) + 2\sum_{i=1}^{N}\omega_i\gamma(d_{i,0}). \tag{12}$$

By taking partial derivatives of $\phi(\omega_i, \mu)$ with respect to ω_i and μ, and by setting the results equal to zero, we can obtain the following equations:

$$\begin{aligned}\frac{\partial\phi(\omega_i,\mu)}{\partial\omega_i} &= -2\sum_{j=1}^{N}\omega_j\gamma(d_{i,j}) - 2\mu + 2\gamma(d_{i,0}) \\ &= 0, \end{aligned} \tag{13}$$

$$\begin{aligned}\frac{\partial\phi(\omega_i,\mu)}{\partial\mu} &= -2\left(\sum_{i=1}^{N}\omega_i - 1\right) \\ &= 0. \end{aligned} \tag{14}$$

Next, by organizing these equations, we can get

$$\sum_{j=1}^{N}\omega_j\gamma(d_{i,j}) + \mu = \gamma(d_{i,0}), \tag{15}$$

$$\sum_{i=1}^{N}\omega_i = 1. \tag{16}$$

Then, N equations related to Eq. (15) can be derived from partial derivatives of $\phi(\omega_i, \mu)$ with respect to all ω_i. Therefore, if we concatenate these results and Eq. (16) and write them in a matrix form, we get the following $N+1$ simultaneous linear equations;

$$\begin{pmatrix} \gamma(d_{1,1}) & \cdots & \gamma(d_{1,N}) & 1 \\ \gamma(d_{2,1}) & \cdots & \gamma(d_{2,N}) & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \gamma(d_{N,1}) & \cdots & \gamma(d_{N,N}) & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix}\begin{pmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_N \\ \mu \end{pmatrix} = \begin{pmatrix} \gamma(d_{1,0}) \\ \gamma(d_{2,0}) \\ \vdots \\ \gamma(d_{N,0}) \\ 1 \end{pmatrix}. \tag{17}$$

Using the above simultaneous linear equations, the optimal ω_i can be derived; we solve Eq. (17) for $(\omega_1, \omega_2, \ldots, \omega_N, \mu)^\top$ using the Gauss-Jordan elimination [44]. Note that $\gamma(d_{i,j})$ in Eq. (17) is calculated from Eq. (9).

E. Re-Modeling of $\overline{P}(\boldsymbol{x}, f_0)$ Using $\boldsymbol{y}_{f_0}$ (if $N_0 > 0$)

Usually, η_d is almost constant regardless of frequency. Thus, if N is sufficiently large, the dependency of the communication distance on f_0 can be accurately estimated using Eq.(6), even if the number of measurement frequencies, M, is small. Meanwhile, although the path loss depends on the frequency in proportion to $\log_{10} f$, if M is small and there is no dataset on f_0, $\overline{P}(\boldsymbol{x}, f_0)$ will be overestimated or underestimated across the area. To improve the path loss modeling over f_0, we put an option in this phase.

This phase uses the optional dataset $\boldsymbol{y}_{f_0}$ that contains a few (i.e., $N_0 \ll N$) measurement values over f_0. Using $\hat{\eta}_d$ estimated in IV-A, we assume $\overline{P}_{\mathrm{est}}(\boldsymbol{x}, f_0)$ as

$$\overline{P}_{\mathrm{est}}(\boldsymbol{x}, f_0) = P'_{\mathrm{C}}(f_0) - 10\hat{\eta}_d\log_{10}||\boldsymbol{x}_{\mathrm{Tx}} - \boldsymbol{x}||, \tag{18}$$

where $P'_{\mathrm{C}}(f_0)$ is a constant value in f_0, which contains both P_C and the effect of frequency. Moreover, P'_{C} can be estimated by fitting $\overline{P}_{\mathrm{est}}(\boldsymbol{x}, f_0)$ into $\boldsymbol{y}_{f_0}$. This can be implemented as follows:

$$P'_{\mathrm{C}}(f_0) = \frac{1}{N_0}\sum_{i=1}^{N_0}\left(P(\boldsymbol{x}_{0,i}, f_0) + 10\hat{\eta}_d\log_{10}||\boldsymbol{x}_{\mathrm{Tx}} - \boldsymbol{x}_{0,i}||\right). \tag{19}$$

We used an updated $\overline{P}_{\mathrm{est}}$ at the end of IV-F.

F. Estimation of Received Signal Power

Finally, the received target value can be estimated using the following equation:

$$\hat{P}(\boldsymbol{x}_0, f_0) = \overline{P}_{\mathrm{est}}(\boldsymbol{x}_0, f_0) + \hat{W}(\boldsymbol{x}_0, f_0). \tag{20}$$

Here, $\overline{P}_{\mathrm{est}}(\boldsymbol{x}_0, f_0)$ is the output of Eq. (6) (or Eq. (18) if $\boldsymbol{y}_{f_0}$ is available). Additionally, $\hat{W}(\boldsymbol{x}_0, f_0)$ is obtained from Eq. (10).

G. Technical Limitations of This Method

The proposed method depends on some radio propagation characteristics such as the shadowing correlation; thus, this fact raises the following fundamental limitations from the viewpoints of applicability;

- Because the proposed method assumes time-invariant shadowing, we may have to periodically update the radio map when it is applied for a time-variant environment (e.g., indoor environment [6]).
- This method is based on the principle that shadowing exhibits a frequency correlation, as mentioned in Sect. III (i.e., related works summarized in Table I). Such a correlation does not necessarily exist in other settings. For example, [45] reported the frequency correlation of approximately 0.4 between 4 GHz and 28 GHz in an outdoor environment; similarly, [46] revealed that the correlation between 980 MHz and 2400 MHz was 0.4 in the curved subway tunnels. Our proposed method utilizes the frequency correlation and thus may not be able to perform an accurate radio map construction in such a situation.
- The number of transmitters in the measurement area should be one because ordinary Kriging assumes weak stationarity for shadowing; i.e., both the variance and correlation

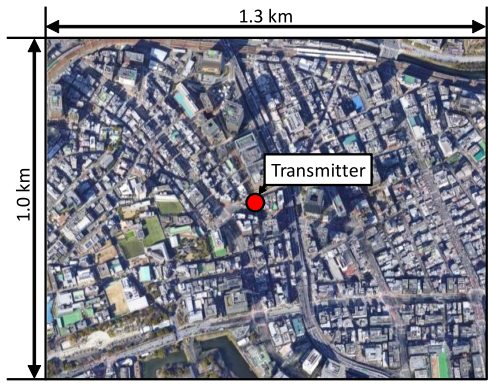


Fig. 2. Aerial photograph of the measurement area (obtained from Google Earth).

distance are invariant over the spatial domain. If multiple transmitters simultaneously transmit signals, this stationarity is broken owing to the inter-transmitter interference, and Kriging may not perform optimal interpolation.

The proposed method will work most effectively in an outdoor environment with a large coverage system (e.g., cellular and television).

One may question whether this method works on both line-of-sight (LOS) and non-LOS (NLOS) situations. Kriging used in the proposed method achieves an optimal estimation for the samples following correlated Gaussian distribution. It is known that the average received power values (in dBm domain) satisfy this condition for both LOS and NLOS (e.g., LOS case in [47], and NLOS case in [43]). Additionally, several works demonstrated the Kriging-based spatial interpolation over LOS [36] and NLOS [48] conditions; empirically, Kriging works in both of them.

V. Measurement Campaign

To evaluate the performance of the proposed method, we used measurements of datasets obtained over cellular bands. The data were measured at Kudanshita Chiyoda-ku in Tokyo, Japan, which can be classified as a typical urban area. This section presents the details of this measurement campaign and fundamental characteristics of the datasets.

A. Measurement Configuration

An aerial photo of this measurement area is shown in Fig. 2. For this measurement, a transmitter with an antenna 17.5 m high was fixed at the center of the measurement area and it emitted continuous waves (CWs) over 870, 2115, and 3500 MHz. All signals were transmitted via an omnidirectional antenna with a transmission power of 29 dBm.

We measured the received signal power values of these bands in an area of 1.0 km $\times$ 1.3 km centered on the transmitter. The reception antenna, which was fixed to the cart, was 1.5 m high. The signal was measured by a spectrum analyzer (Anritsu, MS2712E) with zero-span mode while moving the antenna at approximately 4 km/h and we recorded the received time, position, and signal power at intervals of 0.15 m/s. The major measurement conditions are summarized in Table II.

TABLE II
Measurement Conditions

Transmitter	
Modulation format	CW
Frequency	870, 2115, and 3500 MHz
Antenna	Omnidirectional antenna
Antenna height	17.5 m
Transmission power P_{Tx}	29 dBm
Receiver	
Model number of spectrum analyzer	Anritsu, MS2712E
Measurement mode	Zero span measurements
Resolution band width (RBW)	300 Hz
Antenna	Omnidirectional antenna
Antenna height	1.5 m
Walking speed during measurements	4 km/h
Measurement interval	0.15 m/sec

TABLE III
Frequency Correlation of Shadowing

	870 MHz	2115 MHz	3500 MHz
870 MHz	1.000	0.791	0.780
2115 MHz	0.791	1.000	0.870
3500 MHz	0.780	0.870	1.000

After measurement, we averaged the instantaneous measured values per 10 m grid to eliminate the effect of multipath fading.[3] The numbers of averaged datasets in 870, 2115, and 3500 MHz were 2444, 2445, and 2069, respectively.

B. Fundamental Characteristics of Datasets

We have summarized the fundamental characteristics of the datasets. As shown in Eq. (6), our proposed method assumes that the path loss index η_d is constant, regardless of the frequency and effect of the increase in frequency in proportion to $\eta_f \log_{10} f$. To verify the validity of this assumption, we analyzed the dataset under the following conditions: (i) the dataset was independently analyzed for each frequency with OLS, and (ii) the dataset was analyzed for all frequencies simultaneously assuming Eq. (3).

In the first condition, the estimated η_d, P'_{C}, σ, and d_{cor} are indicated in Table IV; these values were calculated by assuming Eq. (18) as $\overline{P}(\boldsymbol{x}, f)$ and independently processing the dataset for each frequency.

All frequencies showed a similar η_d; however, P'_{C} degraded with an increase in f; this characteristic agrees with Eq. (6). Here, the estimated value of P'_{C} at 870 MHz exceeded the transmission power of 29 dBm. This is because we did not consider the effect of *the reference distance*. Generally, the path loss follows a free-space model within the reference distance; thus, the path loss often follows a piecewise curve [42]. In urban

[3]The grid size should be sufficiently larger than the wavelength (roughly 34 cm at 870 MHz); in contrast, should be smaller than the shadowing correlation distance (dozens to hundreds of meters in practice. For example, 44.24 m at 3500 MHz in our dataset). In other words, it will be preferable to use a size of several meters to several tens of meters. Because the Geospatial Information Authority of Japan (GSI) [49] publishes map information with 10 m grids, we implemented the proposed method with 10 m grids.

TABLE IV
RADIO PROPAGATION CHARACTERISTICS

(The case where the dataset was processed at each frequency separately)				
	σ [dB]	$d_{\rm cor}$ [m]	$P'_{\rm C}$ [dBm]	η_d
870 MHz	8.544	82.68	30.52	4.704
2115 MHz	8.390	47.73	22.50	4.625
3500 MHz	8.456	44.24	12.46	4.484
(The case where all datasets were processed simultaneously)				
	$L_{\rm C}$ [dB]	η_f	η_d	
870-3500 MHz	-58.14	20.00	4.603	

areas, the reference distance is one or ten meters. We do not need to consider this effect because most communication distances of our datasets exceed these values. Additionally, we summarized the path loss characteristics when the dataset was analyzed for all the frequencies simultaneously. This evaluation assumed that the path loss follows Eq. (3), and we extracted $P_{\rm C}$, η_f, and η_d using two-dimensional OLS (Eq. (5)).

Next, we evaluated the frequency correlation of shadowing. We assumed that $\overline{P}(\boldsymbol{x}, f)$ of each frequency follows the parameters in Table IV, and we extracted the shadowing value in each location by calculating $W(\boldsymbol{x}_{k,i}, f_k) = P(\boldsymbol{x}_{k,i}, f_k) - \overline{P}(\boldsymbol{x}_{k,i}, f_k)$. Then, to calculate the shadowing correlation between f_k and f_l, all data at the same grid but different frequencies were used; this operation can be expressed as follows:

$$\rho\left(W(\boldsymbol{x}, f_k), W(\boldsymbol{x}, f_l)\right) = \frac{\sigma_{f_k f_l}}{\sigma_{f_k}\sigma_{f_l}} \approx \frac{\sum_{i=1}^{N_{f_k f_l}} \Delta W(\boldsymbol{x}_{i,f_k f_l}, f_k)\Delta W(\boldsymbol{x}_{i,f_k f_l}, f_l)}{\sigma_{f_k}\sigma_{f_l}}, \tag{21}$$

where

$$\Delta W\left(\boldsymbol{x}_{i,f_k f_l}, f_k\right) = W\left(\boldsymbol{x}_{i,f_k f_l}, f_k\right) - \mathrm{E}\left[W(\boldsymbol{x}_{i,f_k f_l}, f_k)\right],$$
$$\Delta W\left(\boldsymbol{x}_{i,f_k f_l}, f_l\right) = W\left(\boldsymbol{x}_{i,f_k f_l}, f_l\right) - \mathrm{E}\left[W(\boldsymbol{x}_{i,f_k f_l}, f_l)\right].$$

Note that Eq. (21) uses the approximate expression because the expectation values need to be calculated from a finite number of samples.

In Eq. (21), $\sigma_{f_k f_l}$ is the covariance of shadowing between f_k and f_l, and $\boldsymbol{x}_{i,f_k f_l}$ $(i = 1, 2, \ldots, N_{f_k f_l})$ is the location of the i-th grid with the measurement data at both f_k and f_l. Additionally, σ_{f_k} and σ_{f_l} are the standard deviations of shadowing for σ_{f_k} and σ_{f_l}, respectively; the values were obtained from Table IV.

Frequency correlations of shadowing are presented in Table III. Correlations higher than or equal to 0.78 can be observed for any frequency difference. Moreover, because the values of σ are same regardless of f, shadowing values follow a similar probabilistic distribution over all frequencies.

Note that the results obtained in this section can be applied for simulating the radio map construction over space-frequency domains. Our discussion is based on the received power defined in Eqs. (2) and (3). Therefore, this simulation can be implemented using the following two steps: (i) calculate the sum of transmission power and path loss using Eq. (3), and (ii) generate shadowing values from multivariate log-normal distributions. Thus, after the measurement locations and frequencies are determined, we can generate the average received signal values $\overline{P}(\boldsymbol{x}, f)$ using Eq. (3) with $L_{\rm C}$, η_f, and η_d summarized in Table IV. Next, the radio map can be generated after the shadowing value in each measurement point is generated. This shadowing can be generated from the multivariate log-normal distribution, which follows zero-median, and the variance-covariance matrix, which satisfies the frequency correlation matrix shown in Table III, standard deviation σ in Table IV, and correlation distance $d_{\rm cor}$ shown in Table IV.

VI. PERFORMANCE EVALUATION

This section presents the performance of the proposed method using the actual datasets. The performances are compared with those of some related methods.

A. Comparison Methods

We also evaluated the performances of related techniques. Each detail is summarized below.

1) FFNN-Based Path Loss Modeling: This method has been actively discussed in recent years [29], [34], [35] and is recognized as a state-of-the-art technique in path loss modeling.

An FFNN should be designed based on the available information and features of the dataset. For appropriate comparison, we designed the FFNN for our dataset based on related discussions in [29], [34], [35]. Generally, we can easily gain the expression capability of NNs by employing a deep-layered structure [50]; however, related studies have indicated that few (1–3) hidden layers can achieve sufficient accuracy for path loss modeling [29], [33]. Based on these discussions, we exploited one hidden layer. Thus, the FFNN in this study comprises one input layer, one hidden layer, and one output layer. Each neuron is activated by a rectified linear unit (ReLU) function. According to our measurement configuration, the input layer is configured by the following four-dimensional information: communication distance [m], horizontal angle between the transmitter and receiver [rad], and frequency [MHz]. Additionally, the output layer comprises the propagation loss [dB]. Each input or output value was normalized to [0, 1] in each dimension.

The FFNN was trained using both the dataset $\boldsymbol{y}$ and optional dataset $\boldsymbol{y}_{f_0}$ using Adam optimizer [51]. Adam is a first-order gradient-based algorithm for stochastic optimization problems based on adaptive estimates of lower-order moments. This optimizer adaptively scales its learning rates and can avoid locally optimal results in the training phase; empirically, Adam optimizer finds more optimal results than other optimization algorithms such as stochastic gradient descent (SGD) and momentum SGD [52].

In the training phase, we set the initial learning rate and batch size to 0.01 and 64, respectively. Additionally, the number of neurons and epochs were determined using grid-search-based analysis so that the median of RMSE performance where $N = 256$ and $N_0 = 16$ in the interpolation case is minimized; finally, the number of neurons and epochs were designed to be 16 and 300, respectively (the procedure of the RMSE analysis is detailed in VI-C).

2) Empirical Path Loss Model With Parameter Tuning: We also implemented a close-in (CI) free space reference distance model [28] as a related technique. This model aims to develop

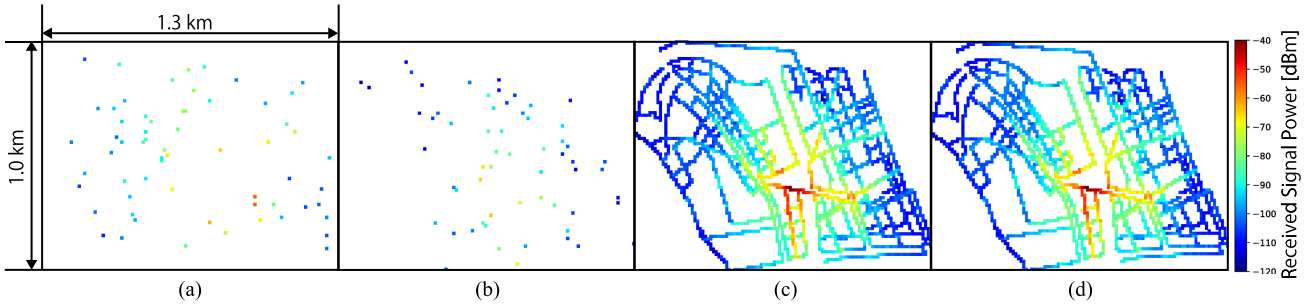


Fig. 3. Example of radio map construction where $N = 128$ and $N_0 = 16$. (a) Dataset (870 MHz). (b) Dataset (3500 MHz). (c) True radio map (2115 MHz). (d) Interpolated radio map (2115 MHz).

an accurate path loss model for 5G urban micro- and macro-cellular systems. This model has one parameter that has to be estimated using the measurement dataset; however, it can accurately estimate the path loss owing to its flexibility. The authors of [28] revealed that this model can estimate the propagation loss accurately in non-line-of-sight at frequency 2-73.5 GHz and communication distance 45-1429 m.

This model assumes the sum of transmission power and the path loss using the following equation:

$$\overline{P}(d, f) = P_{\mathrm{Tx}} - \mathrm{FSPL}(f, 1[\mathrm{m}]) - 10\eta_{\mathrm{CI}} \log_{10} d, \tag{22}$$

where $\mathrm{FSPL}(f, 1[\mathrm{m}])$ is the free-space path loss at $d = 1$ [m]; i.e.,

$$\mathrm{FSPL}(f, 1[\mathrm{m}]) = 20 \log_{10} \left(\frac{4\pi \left(f \times 10^6 \right)}{c} \right). \tag{23}$$

Additionally, c is the speed of light [m/sec], and η_{CI} is the path loss index. This model tunes η_{CI} using the actual dataset via OLS. We calculated this model using the optional dataset $\boldsymbol{y}_{f_0}$, where $N_0 = 16$.

3) Conventional Spatial Interpolation With Kriging: This method assumes that all N samples are selected from the ones in the frequency of interest; i.e., the performance of this method indicates the RMSE performance when we can directly measure sufficient data over f_0. For example, if we construct the radio map over 2115 MHz, this method obtains N samples from 2115 MHz; however, does not use any data on 870 MHz and 3500 MHz.

B. Example of Radio Map Construction

An example of the proposed method is presented in Fig. 3. We attempted to construct a radio map over 2115 MHz from observations over 870 MHz and 3500 MHz. To generate a dataset $\boldsymbol{y}$, we randomly selected a frequency from the dataset band of 870 MHz to 3.5 GHz $N = 128$. Moreover, 16 samples were chosen from the 2115 MHz-band dataset in a random manner to evaluate the target frequency dataset $\boldsymbol{y}_{f_0}$. The constructed radio map expresses the features of both shadowing and path loss.

C. Procedure of RMSE Analysis

To compare the performance accuracies, we evaluated the RMSE via cross-validation-based analysis. This analysis constructed a radio map using the averaged datasets, and the non-gridded (i.e., instantaneous) values were treated as the test dataset to evaluate the accuracy. We performed this evaluation under two situations: (i) frequency interpolation, in which the radio map over 2115 MHz was constructed using observations over 870 MHz and 3500 MHz, and (ii) frequency extrapolation, in which the radio map over 3500 MHz was constructed from observations over 870 MHz and 2115 MHz.

In the first situation, the following procedure was iterated 10 000 times:

1) N samples are randomly selected from the ones at 870 MHz or 3500 MHz. These samples are defined as $\boldsymbol{y}$.
2) N_0 samples are randomly selected from the ones at 2115 MHz. These samples are defined as $\boldsymbol{y}_{f_0}$.
3) Sixteen samples are randomly selected from the ones at 2115 MHz (the dataset selected in step 2 is excluded). These samples are used for cross-validation.
4) We then estimate the received signal power values in the 16 locations selected from the non-gridded dataset.
5) Thereafter, we calculate the RMSE, which can be defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(\hat{P}(\boldsymbol{x}_{0,i}, f_0) - P(\boldsymbol{x}_{0,i}, f_0) \right)^2}, \tag{24}$$

where n is the number of samples for cross-validation, which is 16 in this case. After the above steps were iterated 10 000 times, we obtained the median for the RMSE values.

The evaluation of extrapolation also followed the same procedure as in the above steps, except that the dataset used was different.

D. Results in Frequency Interpolation

The effects of N in the frequency interpolation are shown in Fig. 4. This figure indicates the median values of RMSE, which were obtained from 10 000-times iterated evaluations. We evaluated the performances of the proposed method and the

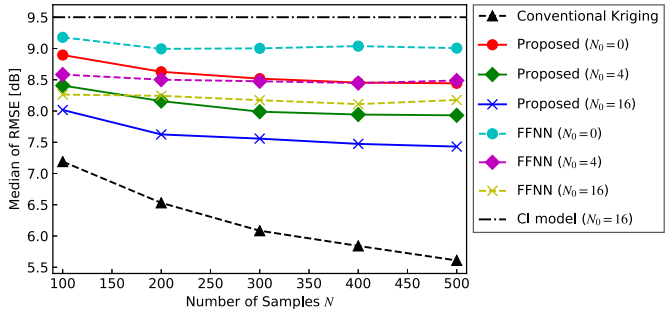


Fig. 4. Median characteristics of RMSE in the case of frequency interpolation (the radio map over 2115 MHz is constructed with N measurement samples over 870 MHz and 3500 MHz and N_0 samples over 2115 MHz). Note that *conventional Kriging* constructs the radio map with N samples over 2115 MHz.

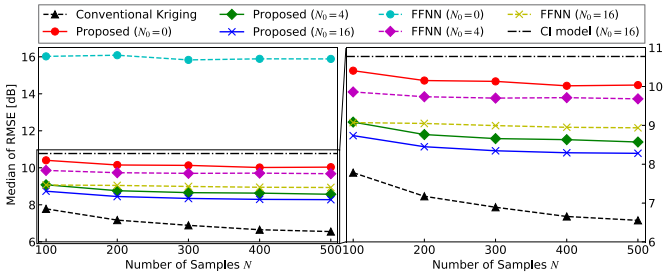


Fig. 5. Median characteristics of RMSE in the case of frequency extrapolation (the radio map over 3500 MHz is constructed with N measurement samples over 870 MHz and 2115 MHz and N_0 samples over 3500 MHz). Note that *conventional Kriging* constructs the radio map with N samples over 3500 MHz.

FFNN-based method, with $N_0 = 0, 4, 16$. Additionally, the CI model assumes $N_0 = 16$; the value of this RMSE is constant regardless of the number of measurement data in $\boldsymbol{y}$ because this model tunes its own parameters using the dataset over only one frequency (i.e., $\boldsymbol{y}_{f_0}$ in this case). The increase in N improves the performances of both the proposed method and the conventional spatial interpolation.

The performance of the proposed method where $N_0 = 16$ is approximately 1-2 dB worse than that of the conventional Kriging model with N samples obtained over 2115 MHz. Meanwhile, the proposed method can improve the RMSE compared with the CI model, even if $\boldsymbol{y}_{f_0}$ is not available. The cause of this difference is that the CI model treats shadowing as an uncertain factor, as shown in Eq. (22). In contrast, the proposed method can estimate both the path loss and shadowing owing to Kriging.

The FFNN-based method outperforms the CI model for any value of N_0 because the FFNN does not assume an explicit equation for the characteristics of the training target. This allows us to deal with complex cases where path loss shows anisotropy, such as direction dependency with reasonable accuracy. In contrast, the CI model assumes that the path loss in all directions and overall areas follows Eq. (22). This assumption degrades the estimation accuracy when the path loss indicates anisotropy; thus, FFNN outperforms the CI model.

However, the proposed method outperforms the FFNN-based method for any value of N_0. In general, the FFNN-based path loss modeling regards shadowing as a Gaussian noise to avoid overfitting [50] and thus cannot predict localities due to shadowing. This fact leads to the advantage of the proposed method in terms of accuracy.

E. Results in Frequency Extrapolation

RMSE performances in the frequency extrapolation are shown in Fig. 5. The relationship between each method is almost equal to that of frequency interpolation; however, if an optional dataset $\boldsymbol{y}_{f_0}$ is not available, the FFNN-based method takes over 15 dB of the RMSE.

If we construct FFNN using the datasets over 870 MHz and 2115 MHz with $N_0 = 0$, its *applicability domain* is between 870 MHz and 2115 MHz. Because FFNN has no information over 3500 MHz, this FFNN takes unpractical RMSE values. In contrast, the proposed method outperforms the FFNN with $N_0 = 0$ and the CI model even though $N_0 = 0$. This is because the proposed method assumes that the path loss is proportional to $\eta_f \log_{10} f$, as shown in Eq. (3). This assumption allows us to roughly estimate the path loss at 3500 MHz; additionally, Kriging can improve the RMSE further.

As with the frequency interpolation situation, the use of $\boldsymbol{y}_{f_0}$ significantly improves both the FFNN and the proposed method. For example, the RMSE value of the proposed method at $N = 500$ is improved to approximately 1.8 dB with $N_0 = 16$.

These results reveal that the proposed method can be used for not only the frequency interpolation but also extrapolation.

VII. Conclusion

We have proposed a powerful method that interpolates the radio map over both space and frequency domains. Experimental evaluations over cellular bands in an urban environment revealed that the proposed method could construct the radio map in frequencies with no (or few) measurement data. The proposed method will improve the performances of radio-map-assisted systems in outdoor environments, such as interference management in cellular systems [10] and fingerprint-based localization systems [9].

Finally, we summarize the future directions the space-frequency-interpolated radio map could adopt. Because the proposed method assumes the shadowing correlation for the frequency domain, its estimation accuracy may be degraded in environments other than those in which frequency correlations have been reported. Related works on the frequency correlation of shadowing have been investigated mainly in outdoor environments, and to the best of our knowledge, no studies have focused on the frequency correlation of shadowing in indoor environments. The extension of the space-frequency interpolation to indoor environments will be an important but challenging direction. To achieve this, we will first need to measure and model such a characteristic. Moreover, shadowing values in indoor environments vary over the time-domain; this fact thus requires real-time updates of radio maps (examples of space-time radio map construction techniques include [5], [53], [54]). To

realize this future direction, we may have to design a radio map construction over time-, space-, and frequency-domains jointly.

Additionally, the use of deep learning techniques, including convolutional neural networks, will improve the accuracy of frequency interpolation in both indoor and outdoor environments. Some recent studies have reported that machine-learning-based radio map construction techniques outperform Kriging in the spatial interpolation over a single frequency [20], [21], [55]. Extension of such a learning technique to the frequency domain will prove to be an interesting future work.

ACKNOWLEDGMENT

The authors would like to thank KDDI Research, Inc. for providing measurement dataset.

REFERENCES

[1] K. Sato, K. Inage, and T. Fujii, "Radio environment map construction with joint space-frequency interpolation," in *Proc. ICAIIC*, Fukuoka, Japan, Feb. 2020, pp. 51–54.

[2] S. Bi, J. Lyu, Z. Ding, and R. Zhang, "Engineering radio maps for wireless resource management," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 133–141, Apr. 2019.

[3] Q. Jiang, Y. Ma, K. Liu, and Z. Dou, "A probabilistic radio map construction scheme for crowdsourcing-based fingerprinting localization," *IEEE Sensors J.*, vol. 16, no. 10, pp. 3764–3774, May 2016.

[4] M. M. Atia, A. Noureldin, and M. J. Korenberg, "Dynamic online-calibrated radio maps for indoor positioning in wireless local area networks," *IEEE Trans. Mobile Comput.*, vol. 12, no. 9, pp. 1774–1787, Sep. 2013.

[5] Y. Tao and L. Zhao, "A novel system for WiFi radio map automatic adaptation and indoor positioning," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 10683–10692, Nov. 2018.

[6] H. Zou, M. Jin, H. Jiang, L. Xie, and C. J. Spanos, "WinIPS: WiFi-based non-intrusive indoor positioning system with online radio map construction and adaptation," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 8118–8130, Dec. 2017.

[7] S. Sorour, Y. Lostanlen, S. Valaee, and K. Majeed, "Joint indoor localization and radio map construction with limited deployment load," *IEEE Trans. Mobile Comput.*, vol. 14, no. 5, pp. 1031–1043, May 2015.

[8] C. Wu, Z. Yang, and C. Xiao, "Automatic radio map adaptation for indoor localization using smartphones," *IEEE Trans. Mobile Comput.*, vol. 17, no. 3, pp. 517–528, Mar. 2018.

[9] Q. D. Vo and P. De, "A survey of fingerprint-based outdoor localization," *IEEE Commun. Surv. Tut.*, vol. 18, no. 1, pp. 491–506, Jan.–Mar. 2016.

[10] J. Perez-Romero *et al.*, "On the use of radio environment maps for interference management in heterogeneous networks," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 184–191, Aug. 2015.

[11] Y. Zeng and X. Xu, "Towards environment-aware 6G communications via channel knowledge map," Jul. 2020. [Online]. Available: https://arxiv.org/abs/2007.09332

[12] X. Mo, Y. Huang, and J. Xu, "Radio-map-based robust positioning optimization for UAV-enabled wireless power transfer," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 179–183, Feb. 2020.

[13] Q. Hu, Y. Cai, A. Liu, and G. Yu, "Joint resource allocation and trajectory optimization for UAV-aided relay networks," in *Proc. IEEE GLOBECOM*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.

[14] R. Murty, R. Chandra, T. Moscibroda, and P. Bahl, "SenseLess: A database-driven white spaces network," *IEEE Trans. Mobile Comput.*, vol. 11, no. 2, pp. 189–203, Feb. 2012.

[15] H. R. Karimi, "Geolocation databases for white space devices in the UHF TV bands: Specification of maximum permitted emission levels," in *IEEE Int. Symp. Dyn. Spectr. Access Netw.*, Aachen, Germany, May 2011, pp. 443–454.

[16] K. Sato, K. Inage, and T. Fujii, "Modeling the kriging-aided spatial spectrum sharing over log-normal channels," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 749–752, Jun. 2019.

[17] K. Sato and T. Fujii, "Kriging-based interference power constraint: Integrated design of the radio environment map and transmission power," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 1, pp. 13–25, Mar. 2017.

[18] C. Phillips, D. Sicker, and D. Grunwald, "Bounding the error of path loss models," in *IEEE Int. Symp. Dyn. Spectr. Access Netw.*, Aachen, Germany, May 2011, pp. 71–82.

[19] N. Cressie, *Statistics for Spatial Data*. Hoboken, NJ, USA: Wiley-Interscience, 1993.

[20] S. I. Rufaida, J. Leu, K. Su, A. Haniz, and J. Takada, "Construction of an indoor radio environment map using gradient boosting decision tree," *Wireless Netw.*, vol. 26, no. 8, pp. 1–22, Jul. 2020.

[21] X. Han, L. Xue, Y. Xu, and Z. Liu, "A two-phase transfer learning-based power spectrum maps reconstruction algorithm for underlay cognitive radio networks," *IEEE Access*, vol. 8, pp. 81232–81245, 2020.

[22] S. Bhattarai *et al.* "An overview of dynamic spectrum sharing: Ongoing initiatives, challenges, and a roadmap for future research," *IEEE Trans. Cogn. Commun. Netw.*, vol. 2, no. 2, pp. 110–128, Jun. 2016.

[23] Y. Okumura, E. Ohmori, T. Kawano, and K. Fukuda, "Field strength and its variability in VHF and UHF land mobile radio service" *Rev. Electr. Commun. Lab.*, vol. 16, pp. 825–873, 1968.

[24] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *IEEE Trans. Veh. Technol.*, vol. 29, no. 3, pp. 317–325, Aug. 1980.

[25] G. Hufford, "ITS irregular terrain model, version 1.2.2, the algorithm," Accessed: Dec. 25, 2020. [Online]. Available: https://www.its.bldrdoc.gov/resources/radio-propagation-software/itm/itm.aspx

[26] D. J. Cichon and T. Kürner, "Digital mobile radio towards future generation systems: Cost 231 final report," COST European Cooperation in the Field of Scientific and Technical Research-Action 231, Tech. Rep., 1993.

[27] C. Phillips, D. Sicker, and D. Grunwald, "A survey of wireless path loss prediction and coverage mapping methods," *IEEE Commun. Surv. Tut.*, vol. 15, no. 1, pp. 255–270, May 2013.

[28] S. Sun *et al.*, "Propagation path loss models for 5G urban micro- and macro-cellular scenarios," in *83rd Veh. Technol. Conf.*, Nanjing, 2016, pp. 1–6.

[29] M. Ayadi, A. B. Zineb, and S. Tabbane, "A UHF path loss model using learning machine for heterogeneous networks," *IEEE Trans. Antennas Propag.*, vol. 65, no. 7, pp. 3675–3683, Jul. 2017.

[30] J. Wen, Y. Zhang, G. Yang, Z. He, and W. Zhang, "Path loss prediction based on machine learning methods for aircraft cabin environments," *IEEE Access*, vol. 7, pp. 159251–159261, 2019.

[31] D. Fernandes *et al.*, "Cloud-based implementation of an automatic coverage estimation methodology for self-organising network," *IEEE Access*, vol. 8, pp. 66456–66474, 2020.

[32] E. Ostlin, H. Zepernick, and H. Suzuki, "Macrocell path-loss prediction using artificial neural networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 2735–2747, Jul. 2010.

[33] S. P. Sotiroudis, S. K. Goudos, K. A. Gotsis, K. Siakavara, and J. N. Sahalos, "Application of a composite differential evolution algorithm in optimal neural network design for propagation path-loss prediction in mobile communication systems," *IEEE Antennas Wireless Propag. Lett.*, vol. 12, pp. 364–367, Mar. 2013.

[34] N. Faruk *et al.*, "Path loss predictions in the VHF and UHF bands within urban environments: Experimental investigation of empirical, heuristics and geospatial models," *IEEE Access*, vol. 7, pp. 77293–77307, 2019.

[35] S. I. Popoola *et al.*, "Determination of neural network parameters for path loss prediction in very high frequency wireless channel," *IEEE Access*, vol. 7, pp. 150462–150483, 2019.

[36] K. Sato, K. Inage, and T. Fujii, "On the performance of neural network residual kriging in radio environment mapping," *IEEE Access*, vol. 7, pp. 94557–94568, 2019.

[37] P. E. Mogensen, P. Eggers, C. Jensen, and J. B. Andersen, "Urban area radio propagation measurements at 955 and 1845 MHz for small and micro cells," in *Proc. IEEE GLOBECOM'91*, Phoenix, AZ, USA, 1991, pp. 1297–1302.

[38] E. Perahia and D. C. Cox, "Shadow fading correlation between uplink, and downlink," in *Proc. IEEE VTS 53rd Veh. Technol. Conf.,* Rhodes, Greece, 2001, pp. 308–312.

[39] B. Van Laethem, F. Quitin, F. Bellens, C. Oestges, and P. De Doncker, "Correlation for multi-frequency propagation in urban environments," *Prog. Electromagn. Res. Lett.*, vol. 29, no. 2, pp. 151–156, 2012.

[40] K. Sato, K. Inage, and T. Fujii, "Frequency correlation of shadowing over TV bands in suburban area," *Electron. Lett.*, vol. 54, no. 1, pp. 6–8, Jan. 2018.

[41] R. Zhang *et al.*, "LTE-unlicensed: The future of spectrum aggregation for cellular networks," *IEEE Wireless Commun.*, vol. 22, no. 3, pp. 150–159, Jun. 2015.

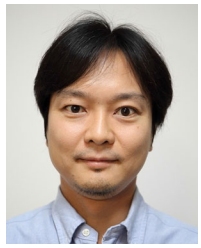
[42] A. J. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[43] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electron. Lett.*, vol. 27, no. 7, pp. 2145–2146, Nov. 1991.
[44] G. Strang, *Linear Algebra and Its Applications*, 2nd ed., New York, NY, USA: Academic Press, Inc., 1980.
[45] D. Burghal, S. L. H. Nguyen, K. Haneda, and A. F. Molisch, "Dual frequency bands shadowing correlation model in a micro-cellular environment," in *Proc. IEEE GLOBECOM*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.
[46] B. Zhang, Z. Zhong, K. Guan, R. He, and C. Briso-Rodríguez, "Shadow fading cross-correlation of multi-frequencies in curved subway tunnels," in *Proc. ITSC*, Qingdao, 2014, pp. 1111–1116.
[47] J. Karedal, N. Czink, A. Paier, F. Tufvesson and A. F. Molisch, "Path loss modeling for vehicle-to-vehicle communications," *IEEE Trans. Veh. Technol.*, vol. 60, no. 1, pp. 323–328, Jan. 2011.
[48] A. Achtzehn, J. Riihijärvi, and P. Mähönen, "Improving accuracy for TVWS geolocation databases: Results from measurement-driven estimation approaches," in *IEEE Int. Symp. Dyn. Spectr. Access Netw.*, McLean, VA, 2014, pp. 392–403.
[49] Geospatial Information Authority of Japan, Accessed: Dec. 25, 2020. [Online]. Available: https://www.gsi.go.jp/ENGLISH/index.html
[50] C. Bishop, *Pattern Recognition and Machine Learning*. Berlin, Germany: Springer, 2006.
[51] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Representations*, San Diego, CA, USA, 2015, pp. 1–41.
[52] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning.* Cambridge, MA, USA: MIT Press, 2016.
[53] S. Kim, E. Dall'Anese, and G. B. Giannakis, "Cooperative spectrum sensing for cognitive radios using Kriged Kalman filtering," *IEEE J. Sel. Top. Signal Process.*, vol. 5, no. 1, pp. 24–36, Feb. 2011.
[54] E. Dall'Anese, S. Kim, and G. B. Giannakis, "Channel gain map tracking via distributed kriging," *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 1205–1211, Mar. 2011.
[55] M. Iwasaki, T. Nishio, M. Morikura, and K. Yamamoto, "Transfer learning-based received power prediction with ray-tracing simulation and small amount of measurement data," in *Proc. IEEE VTC 2020-Fall, Virtual Conf.*, Oct. 2020, pp. 1–6.

Koya Sato (Member, IEEE) received the B.E. degree in electrical engineering from Yamagata University, Yamagata, Japan, in 2013, and the M.E. and Ph.D. degrees from the University of Electro-Communications, Chofu, Japan, in 2015 and 2018, respectively. From 2017 to 2018, he was a Research Fellow (DC2) with Japan Society for the Promotion of Science. He is currently an Assistant Professor with the Tokyo University of Science, Tokyo, Japan. His current research interests include wireless communications, decentralized machine learning, data privacy, and spatial statistics.

Katsuya Suto (Member, IEEE) received the B.Sc. degree in computer engineering from Iwate University, Morioka, Japan, in 2011, and the M.Sc. and Ph.D. degrees in information science from Tohoku University, Sendai, Japan, in 2013 and 2016, respectively. From 2016 to 2018, he was a Postdoctoral Fellow with the Broadband Communications Research Laboratory, University of Waterloo, Waterloo, ON, Canada. He is currently an Assistant Professor with the Department of Computer and Network Engineering, University of Electro-Communications, Tokyo, Japan. His research interests include software-defined networking, mobile edge computing, and Internet of Things. He was the recipient of the Best Paper Award at the IEEE VTC 2013-Spring, the IEEE/CIC ICCC 2015, and the IEEE ICC 2016.

Kei Inage (Member, IEEE) received the B.E. and M.E. degrees in electronic engineering and the Ph.D. degree in communication engineering and informatics from the University of Electro-Communications, Tokyo, Japan, in 2009, 2011, and 2014, respectively. From 2012 to 2014, he was a Research Fellow (DC2) with the Japan Society for the Promotion of Science. From 2014 to 2015, he was a Researcher with the Smart Wireless Laboratory, Wireless Network Research Institute, National Institute of Information and Communications Technology (NICT). From 2015 to 2018, he was an Assistant Professor with Electrical and Electronics Engineering Program, Tokyo Metropolitan College of Industrial Technology, where he is currently an Associate Professor. His current research interests include spectrum sharing and ad hoc networks. He was the recipient of the Young Researcher's Award from the IEICE in 2012.



Koichi Adachi (Senior Member, IEEE) received the B.E., M.E., and Ph.D. degrees in engineering from Keio University, Tokyo, Japan, in 2005, 2007, and 2009, respectively. From 2007 to 2010, he was a Research Fellow with Japan Society for the Promotion of Science. From May 2010 to May 2016, he was with Institute for Infocomm Research, A*STAR, in Singapore. He is currently an Associate Professor with the University of Electro-Communications, Tokyo, Japan. His research interests include cooperative communications and energy efficient communication technologies. He was a Visiting Researcher with the City University of Hong Kong, Hong Kong, in 2009 and the Visiting Research Fellow with the University of Kent, Canterbury, U.K., from Jun. to August 2009. He was the General Co-Chair of the 10th and 11th IEEE Vehicular Technology Society Asia Pacific Wireless Communications Symposium (APWCS), the Track Co-Chair of the transmission technologies and communication theory of the 78th and 80th IEEE Vehicular Technology Conference in 2013 and 2014, respectively, and the Symposium Co-Chair of the communication theory symposium of the IEEE Globecom 2018 and wireless communications symposium of the IEEE Globecom 2020, and the Tutorial Co-Chair of the IEEE ICC 2019. He was an Associate Editor for the *IET Transactions on Communications* between 2015 and 2017, the IEEE Wireless Communications Letters since 2016, the IEEE Transactions On Vehicular Technology between 2016 and 2018, and the IEEE Open Journal Of Vehicular Technology since 2019. He was recognized as an Exemplary Reviewer from the IEEE Communications Letters in 2012 and the IEEE Wireless Communications Letters in 2012, 2013, 2014, and 2015, respectively. He was the recipient of the Excellent Editor Award at the IEEE ComSoc MMTC in 2013. He is a Member of the IEICE.

Takeo Fujii (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in electrical engineering from Keio University, Yokohama, Japan, in 1997, 1999, and 2002, respectively. From 2000 to 2002, he was a Research Associate with the Department of Information and Computer Science, Keio University. From 2002 to 2006, he was an Assistant Professor with the Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology, Fuchu, Japan. From 2006 to 2014, he was an Associate Professor with Advanced Wireless Communication Research Center, University of Electro-Communications, Tokyo, Japan. He is currently a Professor and the Director of the Advanced Wireless and Communication Research Center, University of Electro-Communications. His current research interests include cognitive radio and ad-hoc wireless networks. He is a Fellow of the IEICE. He was the recipient of the Best Paper Award in the IEEE VTC 1999-Fall, the 2001 Active Research Award in Radio Communication Systems from the IEICE Technical Committee of RCS, the 2001 Ericsson Young Scientist Award, the Young Researcher's Award from the IEICE in 2004, the Young Researcher Study Encouragement Award from the IEICE Technical Committee of AN in 2009, the Best Paper Award in the IEEE CCNC 2013, and the IEICE Communication Society Best Paper Award in 2016.